



OIIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/066,007

DATE: 02/14/2002
TIME: 18:39:10

Input Set : N:\Crf3\RULE60\10066007.txt
Output Set: N:\CRF3\02142002\J066007.raw

3 <110> APPLICANT: HOSHINO, Tatsuo
4 OJIMA, Kazuyuki
5 SETOGUCHI, Yutaka
7 <120> TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
9 <130> FILE REFERENCE: ASTAXANTHIN SYNTHETASE
11 <140> CURRENT APPLICATION NUMBER: US/10/066,007
C--> 12 <141> CURRENT FILING DATE: 2001-02-01
14 <150> PRIOR APPLICATION NUMBER: US/09/518,386
15 <151> PRIOR FILING DATE: 2000-03-03
18 <150> PRIOR APPLICATION NUMBER: EP 99104668.1
19 <151> PRIOR FILING DATE: 1999-03-09
21 <150> PRIOR APPLICATION NUMBER: EP 00101666.6
22 <151> PRIOR FILING DATE: 2000-02-01
24 <160> NUMBER OF SEQ ID NOS: 32
26 <170> SOFTWARE: PatentIn Ver. 2.1
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 557
30 <212> TYPE: PRT
31 <213> ORGANISM: Phaffia rhodozyma
33 <220> FEATURE:
34 <221> NAME/KEY: TRANSIT
35 <222> LOCATION: (1)..(26)
37 <400> SEQUENCE: 1
38 Met Phe Ile Leu Val Leu Leu Thr Gly Ala Leu Gly Leu Ala Ala Phe
39 1 5 10 15
41 Ser Trp Ala Ser Ile Ala Phe Phe Ser Leu Tyr Leu Ala Pro Arg Arg
42 20 25 30
44 Ser Ser Leu Tyr Asn Leu Gln Gly Pro Asn His Thr Asn Tyr Phe Thr
45 35 40 45
47 Gly Asn Phe Leu Asp Ile Leu Ser Ala Arg Thr Gly Glu Glu His Ala
48 50 55 60
50 Lys Tyr Arg Glu Lys Tyr Gly Ser Thr Leu Arg Phe Ala Gly Ile Ala
51 65 70 75 80
53 Gly Ala Pro Val Leu Asn Ser Thr Asp Pro Lys Val Phe Asn His Val
54 85 90 95
56 Met Lys Glu Ala Tyr Asp Tyr Pro Lys Pro Gly Met Ala Ala Arg Val
57 100 105 110
59 Leu Arg Ile Ala Thr Gly Asp Gly Val Val Thr Ala Glu Gly Glu Ala
60 115 120 125
62 His Lys Arg His Arg Arg Ile Met Ile Pro Ser Leu Ser Ala Gln Ala
63 130 135 140
65 Val Lys Ser Met Val Pro Ile Phe Leu Glu Lys Gly Met Glu Leu Val
66 145 150 155 160

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/066,007

DATE: 02/14/2002
 TIME: 18:39:10

Input Set : N:\Crif3\RULE60\10066007.txt
 Output Set: N:\CRF3\02142002\J066007.raw

```

68 Asp Lys Met Met Glu Asp Ala Ala Glu Lys Asp Met Ala Val Gly Glu
69          165          170          175
71 Ser Ala Gly Glu Lys Lys Ala Thr Arg Leu Glu Thr Glu Gly Val Asp
72          180          185          190
74 Val Lys Asp Trp Val Gly Arg Ala Thr Leu Asp Val Met Ala Leu Ala
75          195          200          205
77 Gly Phe Asp Tyr Lys Ser Asp Ser Leu Gln Asn Lys Thr Asn Glu Leu
78          210          215          220
80 Tyr Val Ala Phe Val Gly Leu Thr Asp Gly Phe Ala Pro Thr Leu Asp
81 225          230          235          240
83 Ser Phe Lys Ala Ile Met Trp Asp Phe Val Pro Tyr Phe Arg Thr Met
84          245          250          255
86 Lys Arg Arg His Glu Ile Pro Leu Thr Gln Gly Leu Ala Val Ser Arg
87          260          265          270
89 Arg Val Gly Ile Glu Leu Met Glu Gln Lys Lys Gln Ala Val Leu Gly
90          275          280          285
92 Ser Ala Ser Asp Gln Ala Val Asp Lys Lys Asp Val Gln Gly Arg Asp
93          290          295          300
95 Ile Leu Ser Leu Leu Val Arg Ala Asn Ile Ala Ala Asn Leu Pro Glu
96 305          310          315          320
98 Ser Gln Lys Leu Ser Asp Glu Glu Val Leu Ala Gln Ile Ser Asn Leu
99          325          330          335
101 Leu Phe Ala Gly Tyr Glu Thr Ser Ser Thr Val Leu Thr Trp Met Phe
102          340          345          350
104 His Arg Leu Ser Glu Asp Lys Ala Val Gln Asp Lys Leu Arg Glu Glu
105          355          360          365
107 Ile Cys Gln Ile Asp Thr Asp Met Pro Thr Leu Asp Glu Leu Asn Ala
108          370          375          380
110 Leu Pro Tyr Leu Glu Ala Phe Val Lys Glu Ser Leu Arg Leu Asp Pro
111 385          390          395          400
113 Pro Ser Pro Tyr Ala Asn Arg Glu Cys Leu Lys Asp Glu Asp Phe Ile
114          405          410          415
116 Pro Leu Ala Glu Pro Val Ile Gly Arg Asp Gly Ser Val Ile Asn Glu
117          420          425          430
119 Val Arg Ile Thr Lys Gly Thr Met Val Met Leu Pro Leu Phe Asn Ile
120          435          440          445
122 Asn Arg Ser Lys Phe Ile Tyr Gly Glu Asp Ala Glu Glu Phe Arg Pro
123          450          455          460
125 Glu Arg Trp Leu Glu Asp Val Thr Asp Ser Leu Asn Ser Ile Glu Ala
126 465          470          475          480
128 Pro Tyr Gly His Gln Ala Ser Phe Ile Ser Gly Pro Arg Ala Cys Phe
129          485          490          495
131 Gly Trp Arg Phe Ala Val Ala Glu Met Lys Ala Phe Leu Phe Val Thr
132          500          505          510
134 Leu Arg Arg Val Gln Phe Glu Pro Ile Ile Ser His Pro Glu Tyr Glu
135          515          520          525
137 His Ile Thr Leu Ile Ile Ser Arg Pro Arg Ile Val Gly Arg Glu Lys
138          530          535          540
140 Glu Gly Tyr Gln Met Arg Leu Gln Val Lys Pro Val Glu

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/066,007

DATE: 02/14/2002
 TIME: 18:39:10

Input Set : N:\Crf3\RULE60\10066007.txt
 Output Set: N:\CRF3\02142002\J066007.raw

```

141 545
144 <210> SEQ ID NO: 2
145 <211> LENGTH: 1932
146 <212> TYPE: DNA
147 <213> ORGANISM: Phaffia rhodozyma
149 <220> FEATURE:
150 <221> NAME/KEY: CDS
151 <222> LOCATION: (33)..(1706)
153 <220> FEATURE:
154 <221> NAME/KEY: polyA_site
155 <222> LOCATION: (1871)
157 <220> FEATURE:
158 <221> NAME/KEY: mRNA
159 <222> LOCATION: (14)..(1891)
161 <400> SEQUENCE: 2
162 gaattcggca cgaggccacc tactttctcc at atg ttc atc ttg gtc ttg ctc 53
163                                     1 5
164                                     10 15 20
166 aca ggt gct tta ggc ctg gct gct ttc tca tgg gca tcc ata gcg ttc 101
167 Thr Gly Ala Leu Gly Leu Ala Ala Phe Ser Trp Ala Ser Ile Ala Phe
168                                     10 15 20
170 ttc agt ctt tac ctc gct ccg agg cga tct tca ctg tat aac ctt cag 149
171 Phe Ser Leu Tyr Leu Ala Pro Arg Arg Ser Ser Leu Tyr Asn Leu Gln
172                                     25 30 35
174 ggc ccg aat cat acc aac tac ttt aca ggc aat ttt tta gac atc ctc 197
175 Gly Pro Asn His Thr Asn Tyr Phe Thr Gly Asn Phe Leu Asp Ile Leu
176 40                                     45 50 55
178 tca gct cgt aca ggt gaa gag cat gcg aag tac aga gaa aaa tac gga 245
179 Ser Ala Arg Thr Gly Glu Glu His Ala Lys Tyr Arg Glu Lys Tyr Gly
180                                     60 65 70
182 agc acc ctc ccg ttt gct ggg atc gct gga gca ccc gtc ttg aac tcg 293
183 Ser Thr Leu Arg Phe Ala Gly Ile Ala Gly Ala Pro Val Leu Asn Ser
184                                     75 80 85
186 acc gat ccg aaa gtc ttc aac cat gtg atg aaa gaa gcc tac gac tat 341
187 Thr Asp Pro Lys Val Phe Asn His Val Met Lys Glu Ala Tyr Asp Tyr
188                                     90 95 100
190 ccg aaa cct ggt atg gcc gct cga gtg ctc aga att gct acc gga gat 389
191 Pro Lys Pro Gly Met Ala Ala Arg Val Leu Arg Ile Ala Thr Gly Asp
192 105                                     110 115
194 ggt gtt gtt acg gcg gaa ggt gaa gct cat aag cga cat cga agg atc 437
195 Gly Val Val Thr Ala Glu Gly Glu Ala His Lys Arg His Arg Arg Ile
196 120                                     125 130 135
198 atg atc ccc tct ctg tcc gct cag gcc gtt aag tcg atg gtc cca att 485
199 Met Ile Pro Ser Leu Ser Ala Gln Ala Val Lys Ser Met Val Pro Ile
200                                     140 145 150
202 ttc tta gaa aaa ggt atg gaa ctt gtc gac aag atg atg gag gat gcg 533
203 Phe Leu Glu Lys Gly Met Glu Leu Val Asp Lys Met Met Glu Asp Ala
204                                     155 160 165
206 gct gag aag gat atg gcc gtg gga gag tcg gcc ggt gaa aag aag gca 581

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/066,007

DATE: 02/14/2002
 TIME: 18:39:10

Input Set : N:\Crf3\RULE60\10066007.txt
 Output Set: N:\CRF3\02142002\J066007.raw

```

207 Ala Glu Lys Asp Met Ala Val Gly Glu Ser Ala Gly Glu Lys Lys Ala
208      170      175      180
210 acc aga ctc gag acc gaa gga gtc gat gta aag gat tgg gtc ggt cga 629
211 Thr Arg Leu Glu Thr Glu Gly Val Asp Val Lys Asp Trp Val Gly Arg
212      185      190      195
214 gct act ctg gac gtc atg gct ctt gca gga ttt gac tat aag agc gac 677
215 Ala Thr Leu Asp Val Met Ala Leu Ala Gly Phe Asp Tyr Lys Ser Asp
216 200      205      210      215
218 tcg ctc cag aac aag acc aat gag ctc tat gtc gct ttt gtc gga ctt 725
219 Ser Leu Gln Asn Lys Thr Asn Glu Leu Tyr Val Ala Phe Val Gly Leu
220      220      225      230
222 acc gat ggg ttt gct cct acc ttg gac tcg ttc aag gct atc atg tgg 773
223 Thr Asp Gly Phe Ala Pro Thr Leu Asp Ser Phe Lys Ala Ile Met Trp
224      235      240      245
226 gat ttt gta cct tac ttc cga act atg aaa cgg aga cat gag ata cct 821
227 Asp Phe Val Pro Tyr Phe Arg Thr Met Lys Arg Arg His Glu Ile Pro
228      250      255      260
230 ttg act caa gga tta gca gtt tcc cga cga gtt ggg atc gag ctt atg 869
231 Leu Thr Gln Gly Leu Ala Val Ser Arg Arg Val Gly Ile Glu Leu Met
232      265      270      275
234 gag caa aag aag cag gcc gtg ctt ggc tca gct tcc gat cag gct gtt 917
235 Glu Gln Lys Lys Gln Ala Val Leu Gly Ser Ala Ser Asp Gln Ala Val
236 280      285      290      295
238 gat aaa aag gat gtt caa ggt cgg gat atc cta agt ctc cta gtg aga 965
239 Asp Lys Lys Asp Val Gln Gly Arg Asp Ile Leu Ser Leu Leu Val Arg
240      300      305      310
242 gca aac atc gcc gcc aac ctg cct gaa tct caa aag ctg tcc gat gag 1013
243 Ala Asn Ile Ala Ala Asn Leu Pro Glu Ser Gln Lys Leu Ser Asp Glu
244      315      320      325
246 gag gta ctc gct cag atc agt aac ctg tta ttt gct gga tat gaa act 1061
247 Glu Val Leu Ala Gln Ile Ser Asn Leu Leu Phe Ala Gly Tyr Glu Thr
248      330      335      340
250 tct tcg aca gtc ttg aca tgg atg ttt cac cga ctc tca gaa gac aaa 1109
251 Ser Ser Thr Val Leu Thr Trp Met Phe His Arg Leu Ser Glu Asp Lys
252      345      350      355
254 gcc gtt cag gat aaa ctt cga gaa gaa att tgt cag atc gac acg gat 1157
255 Ala Val Gln Asp Lys Leu Arg Glu Glu Ile Cys Gln Ile Asp Thr Asp
256 360      365      370      375
258 atg cct acg cta gac gaa ctt aat gcg ttg cct tat ctc gaa gcg ttt 1205
259 Met Pro Thr Leu Asp Glu Leu Asn Ala Leu Pro Tyr Leu Glu Ala Phe
260      380      385      390
262 gtt aag gag tct ctt cgt cta gac cct cct agt ccg tat gct aac cgt 1253
263 Val Lys Glu Ser Leu Arg Leu Asp Pro Pro Ser Pro Tyr Ala Asn Arg
264      395      400      405
266 gaa tgc tta aag gat gaa gac ttc atc cca ctt gcc gag cct gtc att 1301
267 Glu Cys Leu Lys Asp Glu Asp Phe Ile Pro Leu Ala Glu Pro Val Ile
268      410      415      420
270 ggt cga gat ggg tcg gtc atc aac gag gtc cgg atc acg aaa gga acg 1349
271 Gly Arg Asp Gly Ser Val Ile Asn Glu Val Arg Ile Thr Lys Gly Thr

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/066,007

DATE: 02/14/2002

TIME: 18:39:10

Input Set : N:\Crif3\RULE60\10066007.txt
 Output Set: N:\CRF3\02142002\J066007.raw

```

272      425      430      435
274 atg gtc atg ctt ccg ttg ttc aac atc aat cgt tca aag ttc att tat 1397
275 Met Val Met Leu Pro Leu Phe Asn Ile Asn Arg Ser Lys Phe Ile Tyr
276 440      445      450      455
278 gga gaa gat gca gaa gaa ttc aga ccg gag agg tgg ctt gag gac gta 1445
279 Gly Glu Asp Ala Glu Glu Phe Arg Pro Glu Arg Trp Leu Glu Asp Val
280      460      465      470
282 aca gac tcg ctc aac agt att gaa gca ccc tat gga cac cag gcg agc 1493
283 Thr Asp Ser Leu Asn Ser Ile Glu Ala Pro Tyr Gly His Gln Ala Ser
284      475      480      485
286 ttt atc tct gga ccc aga gct tgc ttt ggt tgg cga ttt gct gtc gcc 1541
287 Phe Ile Ser Gly Pro Arg Ala Cys Phe Gly Trp Arg Phe Ala Val Ala
288      490      495      500
290 gag atg aag gcc ttc ttg ttt gtc act ctc cgt cgg gtc cag ttc gag 1589
291 Glu Met Lys Ala Phe Leu Phe Val Thr Leu Arg Arg Val Gln Phe Glu
292      505      510      515
294 ccc atc atc tct cat cca gag tac gag cac atc acc ttg atc att tcc 1637
295 Pro Ile Ile Ser His Pro Glu Tyr Glu His Ile Thr Leu Ile Ile Ser
296 520      525      530      535
298 cgt cct cga atc gtt ggt aga gag aag gag ggg tac cag atg cgt ttg 1685
299 Arg Pro Arg Ile Val Gly Arg Glu Lys Glu Gly Tyr Gln Met Arg Leu
300      540      545      550
302 cag gtc aag ccg gtc gaa tga gttgattctt catatgttaa gagaagttct 1736
303 Gln Val Lys Pro Val Glu
304      555
306 atatctgaga atgtgtgact aggacaatgc cttctttgta tcgatttggt tctcatacc 1796
308 gggcaggcgc tatgacttct acgtcgtcta tcgtcgtctt ggactctctt cttaccctat 1856
310 atattattcc atccgaaaaa aaaaaaaaaa aaaaaaaaaa aaaaagcggc cgctcgagcc 1916
312 ggctcgtgcc gaattc 1932
315 <210> SEQ ID NO: 3
316 <211> LENGTH: 557
317 <212> TYPE: PRT
318 <213> ORGANISM: Phaffia rhodozyma
320 <400> SEQUENCE: 3
321 Met Phe Ile Leu Val Leu Leu Thr Gly Ala Leu Gly Leu Ala Ala Phe
322 1 5 10 15
323 Ser Trp Ala Ser Ile Ala Phe Phe Ser Leu Tyr Leu Ala Pro Arg Arg
324 20 25 30
325 Ser Ser Leu Tyr Asn Leu Gln Gly Pro Asn His Thr Asn Tyr Phe Thr
326 35 40 45
327 Gly Asn Phe Leu Asp Ile Leu Ser Ala Arg Thr Gly Glu Glu His Ala
328 50 55 60
329 Lys Tyr Arg Glu Lys Tyr Gly Ser Thr Leu Arg Phe Ala Gly Ile Ala
330 65 70 75 80
331 Gly Ala Pro Val Leu Asn Ser Thr Asp Pro Lys Val Phe Asn His Val
332 85 90 95
333 Met Lys Glu Ala Tyr Asp Tyr Pro Lys Pro Gly Met Ala Ala Arg Val
334 100 105 110
335 Leu Arg Ile Ala Thr Gly Asp Gly Val Val Thr Ala Glu Gly Glu Ala

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/066,007

DATE: 02/14/2002

TIME: 18:39:11

Input Set : N:\Crf3\RULE60\10066007.txt

Output Set: N:\CRF3\02142002\J066007.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:727 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:770 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:909 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:942 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27